

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/GB05/000218

International filing date: 21 January 2005 (21.01.2005)

Document type: Certified copy of priority document

Document details: Country/Office: GB

Number: 0401524.4

Filing date: 23 January 2004 (23.01.2004)

Date of receipt at the International Bureau: 17 February 2005 (17.02.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



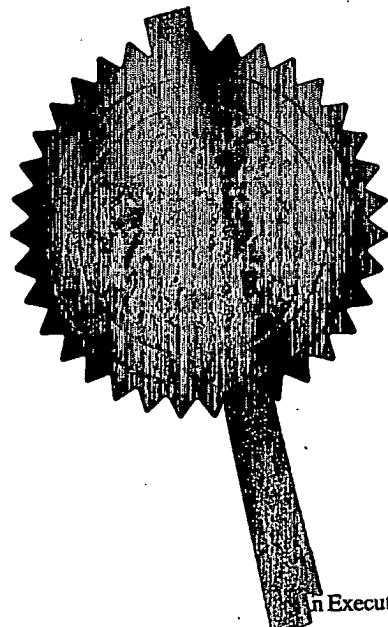
The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

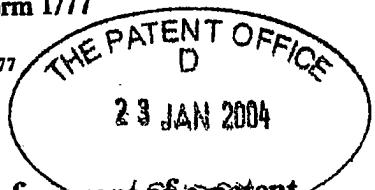
In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



*R. Mahoney*  
Signed   
Dated 10 February 2005



1/77

26JAN04 E0099327 D02890  
P01/7700 0.00-0401524.4 CHEQUE

## Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

 Cardiff Road  
Newport  
South Wales  
NP10 8QQ

1. Your reference

JWJ01015GB

2. Patent application number

(The Patent Office will fill this part in)

0401524.4

23 JAN 2004

3. Full name, address and postcode of the or of each applicant (underline all surnames)

 LingVitae AS  
Trimveien 6, A562  
0372 Oslo  
Norway

Patents ADP number (if you know it)

8612582001

If the applicant is a corporate body, give the country/state of its incorporation

Norway

4. Title of the invention

Method of Analysis

5. Name of your agent (if you have one)

Gill Jennings &amp; Every

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

 Broadgate House  
7 Eldon Street  
London  
EC2M 7LH

Patents ADP number (if you know it)

745002

6. Priority: Complete this section if you are declaring priority from one or more earlier patent applications, filed in the last 12 months.

Country

 Priority application number  
(if you know it)

 Date of filing  
(day / month / year)

7. Divisionals, etc: Complete this section only if this application is a divisional application or resulted from an entitlement dispute (see note f)

 Number of earlier UK application  
(day / month / year)

8. Is a Patents Form 7/77 (Statement of Inventorship and of right to grant of a patent) required in support of this request?

YES

Answer YES if:

- a) any applicant named in part 3 is not an inventor, or
- b) there is an inventor who is not named as an applicant, or
- c) any named applicant is a corporate body.

Otherwise answer NO (See note d)

Patents Form 1/77

9. Accompanying documents: A patent application must include a description of the invention. Not counting duplicates, please enter the number of pages of each item accompanying this form:

Continuation sheets of this form

Description	13
Claim(s)	2
Abstract	
Drawing(s)	4 <i>4</i>

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of Inventorship and right to grant of a patent (Patents Form 7/77)

Request for a preliminary examination and search (Patents Form 9/77)

Request for a substantive examination (Patents Form 10/77)

NO

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

For the applicant

Gill Jennings & Every

Signature

Date 23/01/04

12. Name, daytime telephone number and e-mail address, if any, of person to contact in the United Kingdom

JAPPY, John William Graham

020 7377 1377

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered YES in part 8, a Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) Part 7 should only be completed when a divisional application is being made under section 15(4), or when an application is being made under section 8(3), 12(6) or 37(4) following an entitlement dispute. By completing part 7 you are requesting that this application takes the same filing date as an earlier UK application. If you want the new application to have the same priority date(s) as the earlier UK application, you should also complete part 6 with the priority details.

## METHOD OF ANALYSIS

### Field of the Invention

This invention relates to a method for quantifying the absolute and/or relative numbers of molecules that undergo an analysis procedure; and allows 5 the tracking of an individual molecule during an analysis procedure. The invention is useful especially in the analysis of polynucleotides.

### Background to the Invention

Methods for molecular analysis often require that the original target molecules must be subject to various processes such as amplification and 10 labelling before the analysis itself can take place. It is, however, a problem that the efficiency of such processes are subject to variation. For example, in an amplification process one target molecule in a sample may be copied more times than another target molecule, thereby making it difficult to measure the absolute and relative amounts of the different target molecules that were present in the 15 original sample. Furthermore, the analysis procedure itself often results in the mixing of molecules such that it is not possible to maintain information on each individual molecule. Previously disclosed methods for tagging molecules have not addressed this problem.

Examples of methods of tracking and identifying classes or sub- 20 populations of molecules using oligonucleotide tags have been disclosed in US 5,604,097 and US 5,654,413. US 5,604,097 and US 5,654,413 disclose methods for sorting sub-populations of identical polynucleotides from a sample onto particular solid phase supports. This is achieved by attaching an oligonucleotide tag from a repertoire of tags to each molecule in a population of 25 molecules so that substantially all of the same molecules or same sub-population of molecules have the same tag attached, and substantially all different molecules or different sub-populations of molecules have different oligonucleotide tags attached. Furthermore, each oligonucleotide tag from the repertoire comprises a plurality of sub-units and each sub-unit consists of an 30 oligonucleotide having a length from 3 to 6 nucleotides or from 3 to 6 base pairs; the sub-units being selected to prevent cross-hybridisation. The molecules or sub-populations of molecules may then be sorted by hybridising the

oligonucleotide tags with their respective complements found on the surface of a solid support.

The methods allow tracking and sorting of classes or sub-populations. These methods do not, however, allow tracking of individual target molecules 5 within a class or sub-population of molecules, nor do they allow the quantification of the number of unique molecules involved.

Summary of the Invention

The present invention is based on the realisation that the absolute and/or relative amounts of a unique target molecule can be determined and that 10 individual molecules within a population can be tracked throughout an analysis procedure, by using a molecular tag that is unique to each specific molecule.

According to a first aspect of the invention, a method of quantifying the absolute or relative number of unique molecules present in a sample after carrying out an analysis procedure on the sample, comprises the steps of:

- 15 (i) attaching a unique molecular tag to substantially all of the molecules in the sample;
- (ii) carrying out the analysis procedure using the molecules of the sample; and
- (iii) on the basis of the molecular tags determining the absolute or 20 relative number of unique molecules present in the original sample which underwent the analysis procedure.

The ability to determine the amounts of a unique molecule present in an original sample after amplification is of benefit in many processes. For example, it can be used for transcription analysis in order to measure the amounts of 25 different mRNA classes.

According to a second aspect of the present invention, a method for determining the sequence of a polynucleotide in a sample, comprises the steps of:

- 30 i) attaching a unique molecular tag to substantially all the polynucleotides in the sample;
- ii) fragmenting the amplified polynucleotides; and

iii) sequencing at least those fragmented polynucleotides that comprise a molecular tag, wherein, on the basis of the molecular tags, the sequence information for each individual polynucleotide can be collated.

This is useful in simplifying the reconstruction of sequence data from 5 individual sequence fragments, particularly in *de novo* sequencing.

Description of the Drawings

The invention is described with reference to the accompanying drawings, wherein:

Figure 1 illustrates how the molecular tags are used to identify both the 10 class of molecule and the individual molecule;

Figure 2 illustrates how a further part of the molecular tag can be used to provide sequence information for each molecule;

Figure 3 illustrates how molecules that are attached to substrates such as beads, microbes or cells can be quantified;

15 Figure 4 illustrates how aptamers are used to identify molecules, in solution; and

Figure 5 illustrates the use of two oligonucleotide molecular tags to identify the presence of a polynucleotide sequence.

Detailed Description of the Invention

20 The present invention is used in the analysis of unique molecules. The molecule may be any molecule present in a sample which undergoes an analysis procedure. In a preferred embodiment, the molecules are polymers. The terms "polymer molecules" and "polymers" are used herein to refer to biological molecules made up of a plurality of monomer units. Preferred polymers include 25 proteins (including peptides) and nucleic acid molecules, e.g. DNA, RNA and synthetic analogues thereof, including PNA. The most preferred polymers are polynucleotides.

The term "molecular tag" is used herein to refer to a molecule (or series 30 of molecules) that imparts information about a target molecule to which it is attached. The tag has a unique defined structure or activity that represents the attached individual target molecule. If there are greater than one class of molecules in the sample, the tag may also contain a second defined structure

that represents the class (or sub-population) of target molecule; this part of the tag is termed the "sample identification portion". If the sample comprises a single class of molecules, the sample identification portion is not required and the tag will comprise only the unique portion.

- 5 The molecular tag may be any biological molecule that can impart the necessary information about the target molecule. Preferably, the molecular tag is a polymer molecule that can be designed to have a specific sequence. In the most preferred embodiment, the molecular tag is a polynucleotide that comprises a nucleic acid sequence that is unique and specific for the individual target to
- 10 which the molecular tag is attached. This tag may also comprise a further nucleic acid sequence which is the sample identification portion, that represents the class (or sub-population) of sample molecules. The polynucleotide may be of any suitable sequence. Any suitable size of polynucleotide may be used. The size will depend in part on the number of different target polymers to be "tagged"
- 15 as a unique sequence is required for each (or substantially each) target.

In a further embodiment, the molecular tag is or comprises an aptamer with affinity for the sample molecule. In a preferred embodiment, the molecular tag comprises a target-specific aptamer, (which specifically binds the target molecule) and a unique polynucleotide tag. Aptamers known to recognise

- 20 biomolecules and methods of their production are well known in the art, for example in WO-A-00/71755.

Alternatively, the tag may be a protein. Preferably, the tag in this case is or comprises an antibody which has affinity for the sample molecule.

It is envisaged that a tag could be formed by combining any of the above

- 25 into a single moiety, for example an antibody linked to a polynucleotide or an aptamer linked to a polynucleotide.

Preferably, there is a large excess of unique tags with respect to the sample molecules, such that when attachment occurs it is statistically likely that substantially all sample molecules will be attached to a different, unique tag.

- 30 The sample may comprise molecules that are all identical or substantially similar, or molecules from different populations, i.e. there may be a single class or several classes of molecule in the sample. Molecules in the same class are

identical or have a common attribute, for example a population of identical DNA molecules amplified by PCR, or a mixed population of mRNA transcripts which, although comprising different sequences, all have the common attributes of mRNA and therefore belong to the same class. Molecules of different classes 5 differ in structure or some other attribute, for example a cell surface (as depicted in Figure 3) contains proteins, carbohydrates, glycoprotein, lipids and other biological molecules which all have distinct structures and attributes. Further examples of a sample containing different classes of molecules may be DNA/RNA mixtures, cell lysates, or samples containing different classes of 10 proteins.

It will be apparent to one skilled in the art whether the sample comprises a single class or multiple classes of molecule.

The method of the invention is to be used to "tag" target molecules in a sample prior to analysing the target molecules.

15 Tagging may be carried out by any suitable method, including chemical or enzymic methods, for linking the molecular tag with the target molecule. In the context of a nucleic acid target polymer and a polynucleotide tag, the tagging process may be carried out by suitable ligase enzymes. The tag will usually be ligated onto one of the terminal ends of the target. For example, double 20 stranded polynucleotides may be treated to create single stranded overhangs, which may hybridise with complementary overhangs on the polynucleotide tags and be ligated using a suitable ligase enzyme. Any method of generating the single stranded overhangs may be used, a preferred method is the use of class IIS restriction enzymes.

25 In the context of aptamers or antibodies, the tag is attached to the sample molecule by means of the specific target-aptamer/antibody interaction.

The molecular tag may utilise a binary system, wherein each tag is represented by a series of "0"s and "1"s, allowing a large amount of data to be contained within a small number of tag components. For example, different 30 combinations of "0" and "1" may be formed to provide unique sequences of "0" and "1" that can be used as unique tags.

Preferably, the signals "0" and "1" are represented by different oligonucleotide sequences, for example:

	"0" =	ATTTTTAT
	"1" =	GTTTTTGT
5	ATTTTTATGTTTTGT =	"0,1"
	ATTTTTATATTTTAT =	"0,0"

} unique tags

This system is advantageous since many unique tags can be created using only two units. This is illustrated by Figure 1.

10        When the tag comprises a unique series of "0"s and "1"s according to this binary system, the unique portion of the tag is referred to herein as the "uniqueness number portion". According to the binary system, a preferred tag may comprise a uniqueness number portion, which identifies the individual molecule, and if the sample comprises several classes of molecule, a second 15        defined binary sequence may represent the "sample-identification portion", defining each class of sample molecule. Each class of sample molecule is therefore tagged with a different sample identification portion, and each sample molecule within the class has a different uniqueness number portion. This is illustrated by Figure 1.

20        Attaching the unique portion ("uniqueness number portion" if the binary system is used) of the molecular tag to the sample molecule occurs prior to any analysis procedure. The sample identification portion may be attached to the sample molecule at any point before, during or after the analysis procedure.

25        The analysis procedure may be any procedure used to analyse the molecules.

When the sample molecules are biological molecules such as proteins and polynucleotides, there are a great number of analysis procedures present in the art that would benefit from having each sample molecule individually tagged. Methods of characterising the physical, chemical and functional 30        properties of a molecule are within the scope of "analysis procedures". Such techniques are well known to those in the art. Sequencing of biological polymers may be such an analysis procedure.

The analysis procedure may also comprise the separation of a mixture of molecules, the division of molecules into discrete populations or the amplification of molecules, in particular polynucleotides. These analysis procedures may be applied in many techniques, for example quantifying polynucleotides using the

5 method of the present invention can be used in transcription analysis of cDNA or mRNA, to determine the number of transcripts. Microbial floras may be analysed in a similar fashion; based upon analysis of genomic DNA from different microbial species it is possible to generate unique transcript profiles for each species that can be verified using tags as described by the method of this

10 invention. Quantifying polynucleotides may also be used in ribosomal analysis based on rRNA tagging and detection.

Quantifying molecules that cannot themselves be amplified (as illustrated in Figure 3) may be applied in the analysis of membrane-bound ligands such as proteins, carbohydrates and lipids, and may also be applied in the analysis of

15 biological molecules cross-linked to a surface.

In a preferred embodiment, the analysis procedure comprises amplification by Polymerase Chain Reaction (PCR). Depending on the nature of the molecular tag, only the tag itself or the tag and sample molecule may be amplified.

20 For example, if the tag comprises an antibody attached to a unique polynucleotide, wherein the antibody recognises and binds a protein, amplification by PCR will amplify the unique polynucleotide only. In this embodiment, after contacting the tag to the sample molecule, non-bound tags are removed from the reaction mix. Suitable methods of removal will be

25 apparent to the skilled person. Amplification by PCR is then carried out, wherein only the polynucleotide tag is amplified. The information contained within the tag(s) after amplification is sufficient to determine the number of different molecules present in the original sample.

Alternatively, if both the target molecule and tag are polynucleotides, PCR

30 will result in amplification of both the tag and attached sample molecule. Non-bound tags may again be removed before amplification. In this embodiment, the sample molecules are amplified and may be further analysed or used, whilst the

tags (which have also been amplified) contain the information on the number of different molecules present in the original sample.

In a further preferred embodiment, the analysis procedure comprises detection of the tagged-molecule using a nano-pore detection system. This 5 technique is used when information on each tagged molecule is required. Nanopore methods of detection are well known in the art, and are described in Trends Biotechnol. 2000 Apr; 18(4):147-51.

Suitable nanopores for polynucleotide detection include a protein channel within a lipid bilayer or a "hole" in a thin solid state membrane. 10 Preferably the nanopore has a diameter not much greater than that of a polynucleotide, for example in the range of a few nanometres. As the tagged polynucleotide enters a nanopore in an insulating membrane, the electrical properties of the pore alter. These alterations are measured and as the tagged polynucleotide passes through the pore, a signal is generated for each 15 nucleotide.

The method of the present invention allows an entire sample of polymers to undergo nanopore analysis without losing information on the origin of each molecule, and whilst still being able to determine the number of different molecules present in the original sample, after nanopore analysis.

20 Once the analysis procedure has been carried out, the molecular tags are determined. The method of determination will differ depending on the tag used. When the tag is a polynucleotide, it can be characterised by sequencing. Methods of sequencing are well known to those skilled in the art and suitable techniques will be apparent.

25 Once the sample has been tagged and analysed, it is possible to repeat the method, if required.

The method may be carried out in solution or where the sample molecules are attached to a surface. Such surfaces include biological membranes, beads or living cells. For example, the number of different proteins on a cell surface 30 may be detected, by attaching a unique tag to each class of proteins, amplifying and detecting the number of different unique tags. When the sample molecule is attached to a surface, the molecular tag may comprise an antibody as shown

in Figure 3, although other molecular tags such as aptamers and polynucleotides may also be used.

Figure 3 illustrates a method for quantifying target molecules that are attached to a substrate such as beads, microbes or cells. The method may be 5 used to quantify molecules such as proteins bound to a cell membrane as follows:

- i) The cell is mixed with molecular tags each of which comprises a moiety (antibody or aptamer) with the ability to bind to a specific target molecule, a unique polynucleotide representing the specific target molecule and a sample 10 identification portion. In order to reach saturation of bound target there is a large surplus of molecular tags versus target molecules.
- ii) Any unattached molecular tags are removed from the reaction mix after the binding reaction has reached saturation.
- iii) The polynucleotide part of the molecular tag is amplified and 15 analysed. The number of unique molecular tags that can be associated with a specific target label gives the original number of target molecules.

When the sample molecule is in solution, for example when measuring the number of different mRNA classes in an analysis of transcription, the molecular tag may comprise an aptamer and/or a polynucleotide, as shown in 20 Figure 4, although other molecular tags such as antibodies may also be used.

Figure 4 illustrates quantification of target molecules in solution:

1. Target molecules and molecular tags are mixed. a solution containing the target molecules (e.g. macromolecules such as proteins) is mixed with a large surplus of molecular tags comprising a 25 moiety (e.g. an aptamer) that has the ability to bind to the target molecules with specificity and which comprises a unique polynucleotide portion.
2. Allow molecular tags to bind target molecules.
3. Remove unbound molecular tags. Unbound tags are removed. This can be achieved with gel 30 electrophoresis, spin columns or other separation methods known in the art.
4. Amplify molecular tags bound to target molecules and count the number of unique tags.

The unique tags may then be amplified by PCR before a representative number of the amplified molecular tags are further analysed.

When the sample molecules are polynucleotides, it is possible to use more than one polynucleotide tag in order to increase the specificity of the 5 tagging reaction. Two different tags, each comprising sequences complementary to different but adjacent sequences on the sample polynucleotide and each comprising unique tag sequences, may be hybridised to the sample polynucleotide. These two tags are then ligated together and amplified, as a single polynucleotide, by PCR. The ligation step increases the specificity of the 10 quantification, as two specific tags are required to hybridise compared to the single tag normally used. Only correctly hybridised, adjacent tags will be ligated and amplified. This is illustrated by Figure 5, wherein:

1. Sample polynucleotides and polynucleotide tags are mixed:

Single stranded sample polynucleotides are contacted with two 15 polynucleotide tags each comprising a sequence that can hybridize with specific adjacent parts of the sample sequence. Successful hybridization of the two different polynucleotide tags will bring them into contact with each other, allowing ligation to take place.

2. Polynucleotide tags are hybridised to sample polynucleotides and 20 ligated:

Only the hybridised and ligated polynucleotide tags can be amplified by PCR. The ligation step increases the specificity of the quantification procedure.

3. Polynucleotide tags bound to sample polynucleotides are amplified 25 and the number of unique tags determined.

Figure 1 illustrates a method of the first aspect of this invention wherein the analysis procedure is amplification. The first, pre-amplification sample contains four target polymer molecules, one "A" DNA molecule and three "B" DNA molecules. Prior to the amplification reaction a molecular tag is 30 incorporated onto each target polymer molecule. The molecular tag comprises two portions. One portion is the sample identification portion which identifies the target polymer type. In this example the molecular tag uses a binary system and

subunit "1" represents polymer type "A". Molecular tag subunit "0" represents target polymer type "B". Another portion of the molecular tag, the "uniqueness number portion", identifies the individual target polymer. As can be seen in Figure 1 each of the "B" target DNA molecules has a molecular tag containing 5 a different uniqueness number portion. The molecular tags are incorporated on the targets by ligation.

Once each target polymer molecule has been tagged, the tags and attached targets are amplified using the polymerase reaction. The amplification reaction is random and in any given sample one target polymer molecule may 10 not be copied exactly the same number of times as other target polymer molecules.

After amplification, if a given number of the amplified molecular tags are read, ensuring that each unique molecular tag is read at least once with a high statistical probability, it is possible to deduce the absolute and/or relative amount 15 of "A" and "B" molecules by counting how many unique tags are associated with molecules "A" and "B" respectively.

In this way information is gained about the composition of the first, pre-amplification sample and about the amplification step itself.

A further embodiment of the invention comprises a method of tracking the 20 presence and origin of an individual molecule and/or copies and/or fragments thereof. The sample molecules may be polymeric nucleic acids, which are tagged with oligonucleotide molecular tags as previously described. A preferred analysis procedure is amplification of the tag and attached sample molecule, followed by fragmentation of the amplified polymers; for example as used in "de 25 novo" sequencing methods. The result of this fragmentation is a selection of labelled polynucleotides of different lengths, with all molecules from the same origin (parent molecule) containing the same label, allowing the origin of each molecule to be traced.

The amplified products may be modified in further processes, and the 30 modifications monitored by the incorporation of additional tags. For example, portions of each amplified product may be sequenced.

According to a further aspect of the invention, the sequence of a polynucleotide in a sample may be determined, for example in *de novo* sequencing. This aspect is illustrated by Figure 2.

A molecular tag is attached to substantially all of the polynucleotides in

5 the sample, as described previously. The sample polynucleotides are then fragmented, by methods well known in the art, for example as disclosed in WO-A-00/39333. At least the fragments which comprise a tag may then be sequenced, using methods of polynucleotide sequencing well known in the art. Since there will now be a collection of tagged polynucleotide fragments that,

10 collectively, represent the entire sequence of the original sample molecules, and the origin of each fragment is known due to the tag, re-assembly of the sequence data is simplified.

In a preferred embodiment, the magnifying tag method of sequencing is used, as disclosed in WO-A-00/39333. This describes a method for sequencing

15 polynucleotides by converting the sequence of a target polynucleotide into a second polynucleotide having a defined sequence and positional information contained therein. The sequence information of the target is said to be "magnified" in the second polynucleotide, allowing greater ease of distinguishing between the individual bases on the target molecule. This is achieved using

20 "magnifying tags" which are predetermined nucleic acid sequences. Each of the bases adenine, cytosine, guanine and thymine on the target molecule is represented by an individual magnifying tag, converting the original target sequence into a magnified sequence. Conventional techniques may then be used to determine the order of the magnifying tags, and thereby determining the

25 specific sequence on the target polynucleotide. Each magnifying tag may comprises a label, e.g. a fluorescent label, which may then be identified and used to characterise the magnifying tag. WO-A-00/39333 is incorporated herein by reference.

Another preferred method of sequencing is disclosed by the co-pending

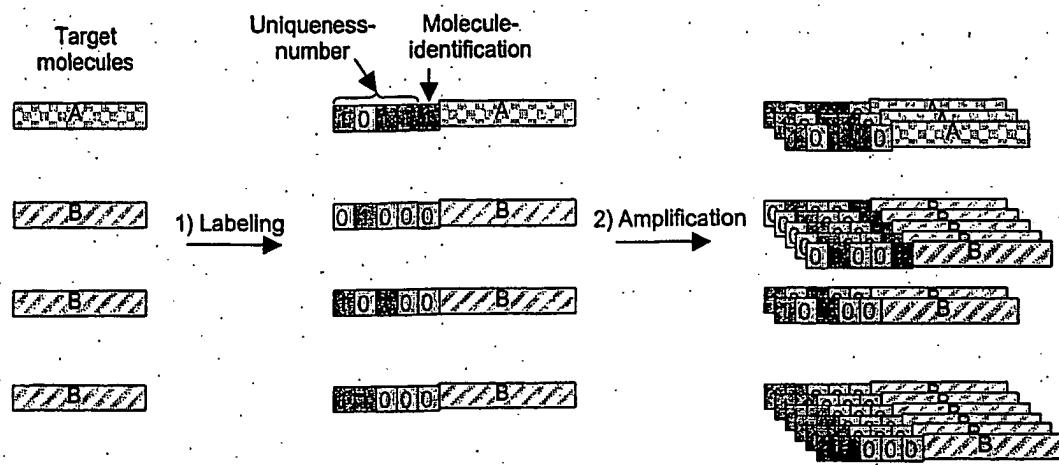
30 patent application GB0308852.3. This is based on the "magnifying tags" method of sequencing, wherein the target polynucleotide sequence is converted into a second "magnified" polynucleotide. The second polynucleotide is then contacted

with at least two of the nucleotides dATP, DTTP, dGTP and DCTP wherein at least one nucleotide comprises a specific detectable label, in order to allow rapid determination of the sequence of the target polynucleotide.

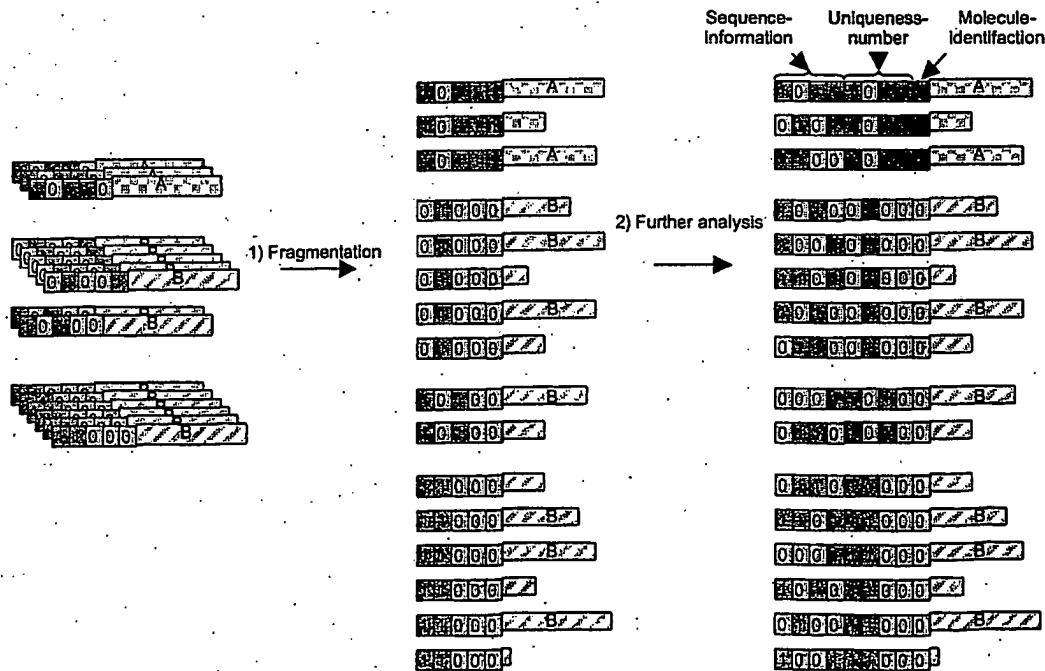
CLAIMS

1. A method of quantifying the absolute or relative number of unique molecules present in a sample after carrying out an analysis procedure on the sample, comprising the steps of:
  - (i) attaching a unique molecular tag to substantially all of the molecules in the sample;
  - (ii) carrying out the analysis procedure using the molecules of the sample; and
  - (iii) on the basis of the molecular tags determining the absolute or relative number of unique molecules present in the original sample which underwent the analysis procedure.
2. A method according to claim 1, further comprising, either before or after analysis, the step of incorporating into the molecular tag a sample identification portion.
3. A method according to claim 1 or 2, wherein the molecules are polymer molecules.
4. A method according to any preceding claim, wherein the sample comprises different molecules.
5. A method according to any preceding claim, wherein the sample comprises multiple molecules of the same type.
6. A method according to any preceding claim, wherein the molecular tag is or comprises an antibody.
7. A method according to any of claims 1 to 5, wherein the molecular tag is or comprises a polynucleotide molecule of defined sequence.
8. A method according to claim 7, wherein the polynucleotide is a DNA molecule of defined sequence.
9. A method according to any of claims 1 to 5, wherein the molecular tag is or comprises an aptamer.
- 30 10. A method according to any of claims 1 to 5, wherein the molecular tags are polynucleotides and the analysis procedure involves an amplification reaction.

11. A method according to claim 10, wherein the molecules are polynucleotides and the analysis procedure involves an amplification of the polynucleotide molecules.
12. A method according to any preceding claim, wherein the analysis procedure involves nano-pore detection.
13. A method according to any preceding claim, wherein the molecular tag, or a part of the molecular tag, indicates the sample-origin of the tagged molecule.
14. A method for determining the sequence of a polynucleotide in a sample, comprising the steps of:
  - i) attaching a unique molecular tag to substantially all the polynucleotides in the sample;
  - ii) amplifying the polynucleotides;
  - iii) fragmenting the amplified polynucleotides; and
  - iv) sequencing at least those fragmented polynucleotides that comprise a molecular tag,wherein, on the basis of the molecular tags, the sequence information for each individual polynucleotide can be collated.
15. A method according to claim 14, wherein the molecular tag is as defined in any of claims 7 to 10 and 13.
16. A method according to claim 14 or claim 15, wherein the sequencing step comprises converting the sequence information into magnifying tags, each tag representing one base in the polynucleotide.

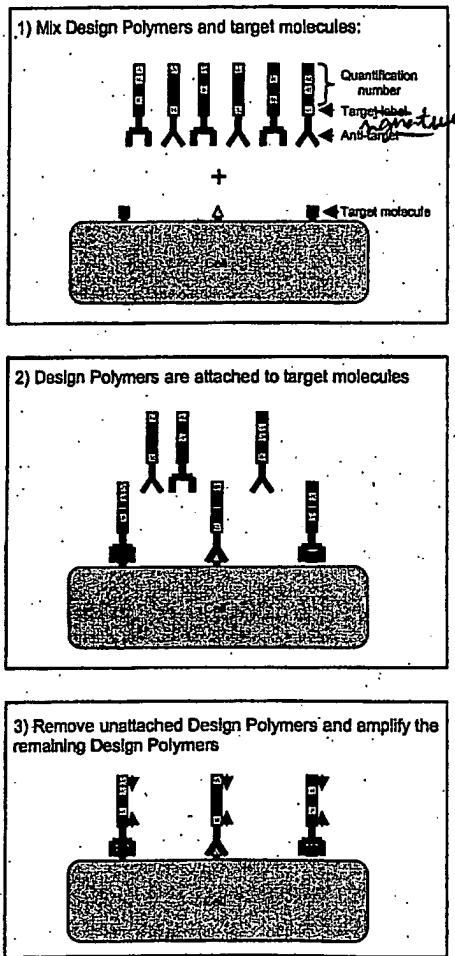


**Figure 1.** The method is illustrated with a solution containing 1 "A" DNA molecule and 3 "B" DNA molecules. The absolute number of target molecules in the starting material and the efficiency of the amplification procedure are normally higher than what is illustrated here. Each target molecule in the solution is tagged by ligation with a Design Polymer representing the molecule ("1" in the first position = "A", "0" in the first position = "B") and four random code units representing the unique tag (1), before they are amplified by PCR (2). If a given number of the amplified Design Polymers are read, ensuring that each unique Design Polymer is read at least once with a high statistical probability, it is possible to deduce the absolute and relative amount of "A" and "B" molecules by counting how many unique tags that are associated with molecule "A" and "B" respectively.



**Figure 2.** *de novo* sequencing. The original molecules are first tagged with the random identification numbers and then amplified (not shown). A fragmentation procedure (for example the erase a base system) are then applied thereby giving rise to target molecules of different lengths. A sequence piece on each end is then converted into a Design Polymer.

Figure 3 – Quantification of membrane bound molecules

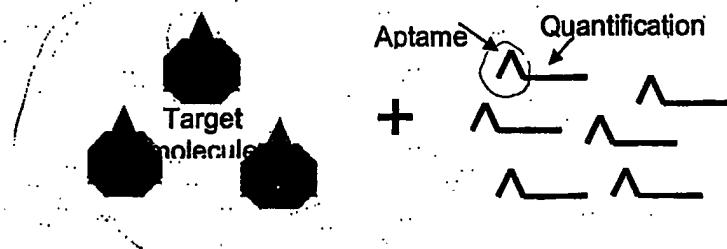


The example relates to methods for quantifying target molecules that are attached to a substrate such as beads, microbes or cells. In this example the method is used to quantify molecules such as proteins bound to a cell membrane.

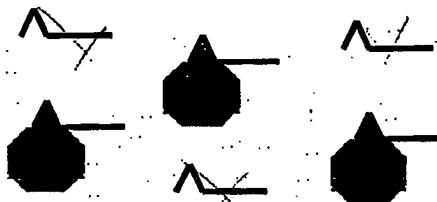
- 1) The cell is mixed with Design Polymer complexes containing a quantification/ uniqueness number, a moiety (antibody, aptamer, etc.) with the ability to bind to a specific target molecule and a target label representing the target that will be bound by the binding moiety. In order to reach saturation of bound target there should be a large surplus of Design Polymers versus target molecules.
- 2) Unattached Design Polymers are removed from the solution after the binding reaction has reached saturation.
- 3) The DNA part of the Design Polymer complex is amplified and analysed. The number of quantification numbers that can be associated with a specific target label gives us the original number of target molecules.

Figure 4 - Quantification of target molecules in solution.

1) Mix target molecules and DesignPolymer complexes:



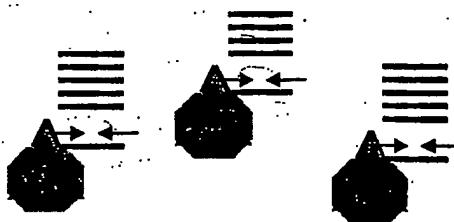
2) Allow DP complexes to bind target molecules:



3) Remove unbound DP complexes:



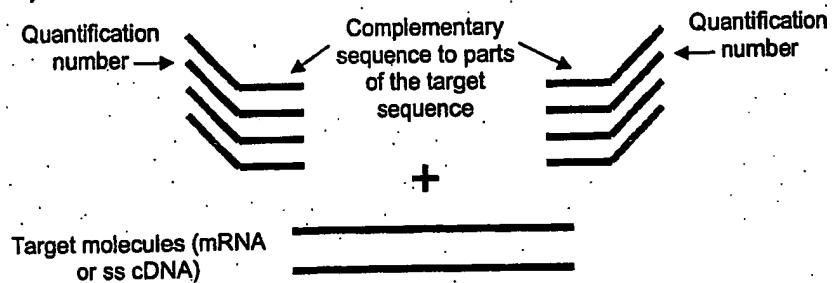
4) Amplify DesignPolymers bound to target molecules and count the number of unique quantification tags:



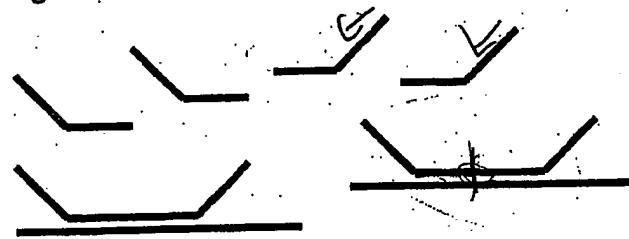
1) a solution containing the target molecules (e.g. macromolecules such as proteins) is mixed with a large surplus of DP complexes (DNA) with one moiety (e.g. an aptamer) that will bind the target molecules with specificity and one moiety representing a random quantification number. 2) and 3) Unbound DP complexes are removed after binding has been allowed. The separation can be achieved with gels, spin columns or other separation methods present in the art. 4) The quantification numbers are then amplified with PCR before a representative number of the amplified quantification numbers are analysed.

Figure 5 - Quantification system combined with a ligation step in order to increase specificity

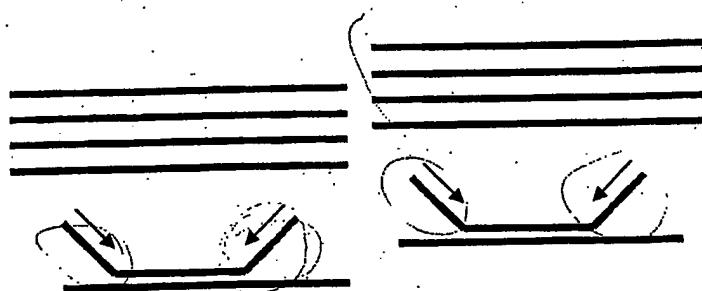
1) Mix mRNA/ cDNA and DesignPolymer complexes:



2) Hybridize DP complexes to target molecules and ligate:



3) Amplify DesignPolymers bound to target molecules and count the number of unique quantification tags:



1) Single stranded target molecules are mixed with DesignPolymer complexes containing a moiety that can hybridize with specific parts of the target sequence. Successful hybridization of the two different DP complexes will bring them into contact with each other, thereby allowing ligation to take place in step 2). Only the bound and ligated DP complexes can be amplified with exponential PCR and the ligation step can thus be used to increase the specificity of the quantification procedure.